DOCUMENT RESUME

ED 277 856 CE 046 089

AUTHOR Kusy, Mitchell E., Jr.
TITLE The Effects of Types of Training Evaluation on Support of Training among Corporate Managers.
PUB DATE Dec 86
NOTE 29p.; Paper presented at the American Vocational Association Convention (Dallas, TX, December 1986). For the report on which this speech was based, see ED 270 640.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Administrator Attitudes; Case Studies; *Employer Attitudes; *Evaluation Methods; *Industrial Training; *On the Job Training; Postsecondary Education; *Program Effectiveness; *Program Evaluation

ABSTRACT

A study was conducted to determine which type of training evaluation method elicited the most management support of the training function among corporate managers. The investigator designed and distributed a case study survey instrument called the Training Evaluation Methods Survey (TEMS) to assess the extent of management support for each type of training evaluation method--reaction, learning, behavior, and results. MBA students with management experience and nontraining managers participated in the study. Data were analyzed using analysis of variance, a chi-square, and descriptive statistics. For both samples, the data indicated that the results evaluation format received the most support, with progressively less support for behavior, learning, and reaction evaluations. The data were significant for all sample groups. Percentages of respondents who gave the results evaluation format the most support for each sample group ranged from 79 percent to 86 percent. There was no significant difference between the MBA group and the manager group. There were no interaction effects between support of each evaluation type and title, years of experience, number of people managed, annual budget, or sex. The study provides training practitioners with a mandate for demonstrating training results to top management and an accountability system for gaining management support of the training function that affects positively the outcome of a business organization. A two-page reference list concludes the document. (Author/KC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED277856

THE EFFECTS OF TYPES OF TRAINING EVALUATION ON SUPPORT OF TRAINING AMONG CORPORATE MANAGERS

BY

MITCHELL E. KUSY, JR.

Biographical Data

Mitchell E. Kusy, Jr. is Manager of Employment and Development at Group Health, Inc. in Minneapolis. In this and previous positions, Dr. Kusy's responsibilities have included human resources development, organization development, employee/labor relations, employment, and AA/EEO. Dr. Kusy has a Bachelor's Degree in Psychology from Providence College, a Master's degree in Psychiatic Social Work from the University of Wisconsin and a Master's degree in Industrial Relations from the University of Minnesota. He holds a Doctorate from the University of Minnesota in Vocational Education with a specialization in Training and Development.

Manager, Employment and Development
Group Health, Inc.
2829 University Ave. SE
Minneapolis, MN 55414

612-623-8516

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M E Kusy Jr.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CE046089

BEST COPY AVAILABLE

ABSTRACT

With organizations becoming leaner during times of cost consciousness, training departments are having to demonstrate their contributions to organizations' bottom lines by becoming more productivity based than they have been historically. For training professionals to obtain the needed financial and organizational support of their training efforts, top management should corroborate the training function. To gain this support, training professionals have evaluation procedures available to them. Evaluation of an organization's training efforts is a primary vehicle for demonstrating training's impact on an organization's bottom line.

The purpose of this study was to determine which type of training evaluation method elicited the most management support of the training function among corporate managers. The investigator designed and distributed a case study survey instrument (TEMS) to assess the extent of management support of training for each type of training evaluation method--

reaction, learning, behavior, and results. MBA students with management experience and non-training managers participated in the study.

Prior to the actual study, the investigator conducted reliability and validity studies. Results indicated that the TEMS was reliable in a test-retest format and valid through a face validity procedure. The data analysis consisted of an ANOVA, a chi-square, and descriptive statistics.

For both samples, the data indicated that the results evaluation format received the most support, with progressively less support for behavior, learning, and reaction evaluations, respectively. The data were significant for all sample groups, with values $p<.0001$ for each group using an ANOVA and values $p<.001$ using a chi-square. Percentages of respondents who gave the results evaluation format the most support for each sample group ranged from 79% to 86%. There was no significant difference between the MBA group and the manager group. There were no interaction effects between support of each evaluation type and title, years of experience, number of people managed, annual budget, or sex.

This study provides training practitioners with a mandate for demonstrating training results to top management and an accountability system for gaining management support of the training function that affects positively the outcome of the organization.

PROBLEM STATEMENT

With organizations becoming leaner during times of cost consciousness, training departments are having to demonstrate their contributions to an organization's bottom line by becoming more productivity based than they have been historically. Of significance is the demonstration of how training contributes to overall organizational productivity. In the past, upper management simply wanted to know that trainers thought training was relevant and that learning occurred. Brown (1980) noted that more recently,

> Upper management has begun to look at all training as an investment. By looking at training as an investment, it has become increasingly important to determine what economic benefits may be directly attributed to training programs. This recent emphasis on return on investment for training dollars has made several evaluation techniques of prime importance for training directors who are being held accountable for the results their programs are designed to achieve (p. 11).

A common complaint of training professionals is lack of management support of training (Dobbs, (1980). Unfortunately, training managers have not consistently conducted effective evaluation programs for convincing top management that training can contribute to an organization's bottom line results. According to a survey of Training and Development Journal readers, one third regarded evaluation as their most difficult task (Galagan, 1983). Historically, management has responded to training's lack of association with an organization's productivity by reducing or eliminating training in those departments that do not have a substantial impact on an organization's bottom line results, particularly during times of budget trimming (Barta, 1982; Kelley, Orgel & Baer, 1984; Zenger & Hargis, 1982). These cost conscious measures have placed pressure on training personnel to demonstrate to those management staff responsible for budgeting operations that training can contribute to the productivity of the organization. The problem becomes one of discerning how to most effectively evaluate training to demonstrate its impact on the organizational bottom line.

REVIEW OF LITERATURE

There are four types of evaluation methods that training personnel typically use to assess training effectiveness. These include the following that Kirkpatrick identified --

1. Reaction by the trainee to training
2. Learning which occurred by the trainee
3. Behavior changes of the trainee on the job
4. Results that reflect changes in organizational or departmental productivity (Brethower & Rummler, 1976; Dopyera & Pitone, 1983; Kirkpatrick, 1975, 1977, 1979, 1983a, 1983b; Swierczek & Carmichael, 1985).

Several researchers have investigated the trainer's use of each of the four evaluation types--reaction, learning, behavior, and results. In one study, researchers found that reaction evaluation formats had the highest frequency of use with progressively less frequency of use for learning, behavior, and results evaluation formats, respectively (Catalanello & Kirpatrick, 1975). In a survey of 285 training officers in industry, training officers stated that they used reaction evaluation

73% of the time, learning 82%, behavior 19%, and results 12% (Smeltzer, 1975).

Managers are demanding greater accountability for training to prove its worth to the organization. "'Accountability' means requiring the training staff to produce documented evidence of training quality and efficiency on a regular or periodic basis (Smith, 1980, p.74). Top management is seeking and demanding proof that the training investment pays off for the organization through improved productivity (Salinger, 1981).

A strategy for determining training's impact on bottom line productivity is use of the results oriented evaluation format with the multiple baseline design. The multiple baseline design assesses the same training program with different groups at staggered points in time. Two types of multiple baseline designs are available to training professionals for evaluation of training efforts:

- Multiple baseline design across behaviors or results variables
- Multiple baseline design across subjects or groups.

The first design, across behaviors or results vaiables, evaluates the effects of one training intervention on two or more different behaviors or productivity results. The second design, across subjects or groups, assesses the effects of a training intervention on two or more classifications or groups of employees.

The multiple baseline design is the recommended approach for evaluating the impact that training has on either behavior or results because it is practical and effective in the business world (Brown, 1980).

Traditionally, most training professionals have not incorporated the results-oriented method into their evaluation strategies. Top management is seeking accountability that the training investment pays off for the organization through improved productivity. What is significant is the fact that the productivity indicators used in results-oriented evaluations demonstrate to top management that training can contribute to the financial well-being of the organization (Trapnell, 1984).

HYPOTHESES

According to the literature, results-oriented evaluations are associated with training accountability, and accountability is what top management demands of the training function. From this perspective, it would appear that corporate managers would be most likely to support those training efforts that use results evaluation. Subsequently, the research question becomes one of determining the accuracy of this conclusion--are corporate managers more likely to support training based on results evaluations than other evaluation methods?

Null Hypotheses:

Ho_1 There is no significant difference in support of training based on evaluation types consisting of (a) reaction, (b) learning, (c) behavior, or (d) results, among MBA managers.

Ho_2 There is no significant difference in support of training based on evaluation types consisting of (a) reaction, (b) learning, (c) behavior, or (d) results, among non-training managers.

Ho_3 There is no significant difference in support of training based on evaluation types consisting of (a) reaction, (b) learning, (c) behavior, or (d) results, among a combined sample of non-training managers and MBA managers.

Directional Hypotheses:

H_1 MBA managers are more likely to support training which uses results evaluation than either reaction, learning, or behavior evaluaton.

H_2 Non-training managers are more likely to support training which uses results evaluation than either reaction, learning, or behavior evaluation.

H_3 Non-training and MBA managers are more likely to support training which uses results evaluation than either reaction, learning, or behavior evaluaiton.

METHODS

Two populations in this study included part-time MBA students and non-training managers in industry. A total of 90 MBA students from four required classes from a local college participated. The investigator eliminated from the study training managers and respondents who were not managers. The investigator did not replace these eliminated individuals with additional subjects.

The second population consisted of a random sample of 75 non-training managers from a total of 150 managers within a health care company headquartered in the Twin Cities.

The investigator used the researcher-developed Training Evaluation Methods Survey (TEMS) to assess the level of management support of training for each type of evaluation method. The TEMS described four scenarios, each of which illustrated one of the four evaluation types--reaction, learning, behavior, and results--that Kirkpatrick described.

To be certain that each scenario description was an accurate illustration of each evaluation type, the

investigator used expert opinion for arriving at the validity of these descriptions. The pilot study consisted of the investigator disseminating the survey to a class of 20 graduate students in the Training and Development Program at a local university.

To establish the reliability of the TEMS, the investigator randomly distributed the TEMS to 20 exempt professionals in three departments at a local health care organization. Using a test-retest reliability procedure, the investigator asked these individuals to complete the TEMS. One week after the first administration of the TEMS, the investigator readministered the TEMS to this same group of individuals.

To reduce the bias associated with the order of presentation of the evaluation scenarios, the investigator randomly ordered these scenarios. The MBA students received the TEMS in their class from the investigator. The investigator instructed the MBA students to complete the TEMS during class time. For the manager sample, the investigator randomly selected 75 out of a total of 150 managers at a local health care organization.

This single factor block design study included four decisions--corresponding to each evaluation scenario--that each subject made regarding level of support. The investigator used a two-way ANOVA and a chi-square to analyze the data. In addition, the investigator conducted a chi-square analysis to determine if there was a significant difference between the MBA group and the manager group.

In order to determine if there were interaction effects that contributed to the data obtained, the investigator collected demographic information on each respondent and graphed scatter plots to assess interaction effects.

RESULTS

Results from the validation procedure indicated that the TEMS measured the four evaluation types--reaction, learning, behavior, and results. Table 1 lists the Pearson product-moment correlation coefficients for each section of the TEMS. Due to the high reliability coefficients for each section, the investigator determined the reliability of the instrument.

In the MBA sample, the response rate was 100%. The investigator eliminated 27 respondents from the study; these respondents were either non-managers or training managers. With the 27 respondents eliminated, there were 63 completed TEMS which the investigator used in the data analysis. In the manager sample, the response rate was 56%.

Scores on the TEMS ranged from 1 through 5 with the following benchmark values:

1 - Unlikely to give support

2 - Minimally likely to give support

3 - Somewhat likely to give support

4 - Likely to give support

5 - Very likely to give support.

Table 2 lists the average support scores and standard deviations for all three subject groups--MBA students, managers, and the combined sample of MBA students and managers. For all three subject groups, the results evaluation format received the most support, with progressively less support for behavior, learning, and reaction evaluations, respectively.

For the MBA sample, the ANOVA and chi-square indicated significant results, $F(3, 186) = 64.05$, $p<.0001$ and χ^2 $(3, N = 63) = 125$, $p<.001$, demonstrating that the results evaluation received the most support. For the manager group, the ANOVA and chi-square indicated significant results, $F(3, 123) = 17.08$, $p<.0001$ and $\chi^2(3, N = 42) = 65.6$, $p<.001$, demonstrating that the results evaluation received the most support.

For the combined group, the ANOVA demonstrated significant results, $F(3, 312) = 74.71$, $p<.0001$.

Table 2 also shows the percentage of respondents who selected the evaluation scenario for which they gave the most support for each subject group. These descriptive statistics clearly demonstrate the high degree of support of the results evaluation format by corporate managers.

The investigator also computed a chi-square to determine if there was a significant difference between the MBA group and the manager group. Results showed a $\chi^2(2, N = 105) = 1.84$, indicating no significant difference between the MBA group and the manager group. There were no interaction effects

between any of the evaluation types and title, years of experience, number of people managed, annual budget, or sex.

CONCLUSIONS

This study provides trainers with a mandate for demonstrating training results to top management and for gaining management support of training.

With financial accountability of increasing concern in industry it is becoming necessary to consider training approaches that are associated with this financial perspective. This research demonstrates an effective approach for establishing accountability through the results-oriented evaluation format.

This research demonstrates similar results for both samples, showing that the model produces consistent data and is not dependent on subject group. Related to this result is an implication for further research. Since this investigator sampled only two population groups, it may be appropriate to test this model on a wide range of industries in

order to demonstrate the generalizability of results across populations. The results from this study are generalizable to two population groups--2nd year MBA students who are non-training managers in required courses at the College of St. Thomas and non-training managers at Group Health, Inc. The current research has set the pace for the replication of an effective model in industries such as high technology, finance, service, health care, retail, and manufacturing.

It may be useful to consider using this model with a much larger number of subjects. The investigator was not able to determine significant interaction effects; this situation may have been due to a relatively small sample size. A larger sample size could contribute to more effective discrimination of interaction effects with the demographic variables used in the study. As Hays (1963) noted, increasing the sample size makes a given statistical test more powerful. Associated with assessing the effects different variables have on the results, future researchers may wish to consider the effects which other variables may have on support of training. Some of these intervening

variables may include manager salary, size of the organization, type of industry, and organizational dollars spent on training.

This study assessed support through selected evaluation strategies associated with each evaluation type. The investigator selected the various evaluation strategies, such as multiple baseline and pre-post formats, based upon what the research literature indicated would be appropriate assessment methods. Of significance is the fact that there were other assessment methods which the investigator could have selected. Subsequently, it may be useful to assess level of support of training through the use of other evaluation strategies. For example, future researchers may wish to assess level of support of the behavior evaluation format through a strategy using a control group design rather than a multiple baseline design. Another possibility is to assess learning through a pre-then-posttest format, as discussed earlier in this paper, rather than a pre-post format. With the development and refinement of a number of assessment strategies, training professionals may have more relevant evaluation tools available.

Related to this issue of using other evaluation strategies, it may be important to assess which evaluation strategy is most effective in obtaining management support within each evaluation method. For example, it may be useful for training professionals to determine, from among several evaluation strategies, which evaluation strategy receives the most support within one evaluation format such as the results evaluation format. This is a refinement of the current research which used only one strategy, the multiple baseline design, for assessing level of management support of the results evaluation format. There could potentially be other designs which are more effective in gaining management support of the training function than the use of the multiple baseline design with the results evaluation format. This method would facilitate the development of potentially more meaningful evaluation strategies for each evaluation method.

The investigator had selected sales figures as the organizational result in the results evaluation format. It would be interesting to assess other organizational results to determine the level of

management support for each of these. Future research could potentially consider such variables as number of grievance claims, accident records, turnover figures, absenteeism data, and scrap rates. Support of these variables could be dependent on such factors as type of industry and the various functional responsibilities of the management group assessed. The implication for further research would be to design a study which determines which type of organizational results receives the most support, with consideration of type of industry and functional management responsibilities of the group surveyed.

While the investigator desired a return rate greater than 56% for the manager sample, the structure of the study reduced the negative effects associated with a lower return rate. Specifically, since the investigator conducted an identical study with the MBA students who provided a 100% response rate and the results were significant for each group, the 56% response rate is not of critical concern. Since there were no significant differences between groups, the lower response rate for the manager group becomes of even less concern.

A few individuals commented on the TEMS that they could see value in combining evaluation methods. This idea is worth exploring in further detail. While the investigator asked respondents to select only one choice related to their overall level of support, it may be that a combination of methods is more effective than one single evaluation method. This study did not assess the effects which combinations of evaluation formats have on level of managerial support. Because of the progression in level of support from reaction to learning to behavior to results, the research suggests that perhaps behavior and results evaluation formats might be an effective combination to consider. This suggestion can only be subject to conjecture at the present time. Further research might validate this supposition.

Another future direction for training professionals to consider is the contribution which results-evaluation data can make in the area of cost-benefit analyses. This is an important consideration in planning results evaluation studies because the data gathered from these studies must be translatable into a cost-benefit formula; this is instrumental in

determining whether the productivity/bottom line results from training yield a significant return on investment. A high level of productivity resulting from training is meaningless if the training program costs more than the productivity results achieved. Subsequently, training professionals need more rigorous training, not only in evaluation strategies, but also in designing cost-benefit analyses. According to Mangum (1984), and Mosier (1986), while training directors acknowledge the significance of cost-benefit analyses of training programs, these analyses usually take the form of informal, unstructured, "gut feelings" computations rather than more formal, data intensive approaches. Therefore, it is imperative that training professionals consider the interaction of training evaluation methods with cost-benefit analysis studies.

Continuing research is necessary to determine the most effective evaluation strategies for demonstrating training accountability. Increased educational opportunities are needed to provide trainers with the prerequisite skills for effectively applying these strategies.

Without a training accountability system such as the one presented here , the vitality of the training function is likely to remain at the discretion of non-training professionals who control the budgeting process. With such an accountability system, training will assume a more proactive and professional stance with greater force to affect positively the outcome of the organization.

Table 1

Test-retest Reliability Coefficients (N = 20) for Each Evaluation Section and Final Overall Section of the TEMS

Section of TEMS	Reliability coefficient
Reaction	.93
Learning	.96
Behavior	.89
Results	.88
Final	.95

Table 2

Average Responses, Standard Deviations, and Percent of Support for Each Type of Training Evaluation Method for Each Subject Group

	Values on a 5 point Likert scale		
Subject group	Average response	Standard deviation	Percent of support
MBA (N = 63)			
Reaction	2.80	.96	0
Learning	3.39	.91	5
Behavior	3.57	.88	9
Results	4.44	.59	86
Manager (N = 42)			
Reaction	3.33	1.00	0
Learning	3.63	1.05	12
Behavior	3.75	.87	9
Results	4.44	.55	79
Combined (N = 105)			
Reaction	3.01	1.02	0
Learning	3.49	.97	8
Behavior	3.64	.87	9
Results	4.44	.5[illegible]	83

REFERENCES

Barta, T. (1982). Methods to determine return on the training investment. Performance and Instruction, 21 (1), 16-17, 24.

Brown, M. G. (1980). Evaluating training via multiple baseline designs. Training and Development Journal, 34 (10), 11-16.

Catalanello, R. F. & Kirkpatrick, D. L. (1975). Evaluating training programs - The state of the art. In D. L. Kirkpatrick (Ed.), Evaluating Training Programs (pp. 258-262). Madison, WI: American Society for Training and Development.

Dobbs, J. H. (1980). Building training department credibility. Training and Development Journal, 34 (3), 14-21.

Galagan, P. (1983). The numbers game: Putting value on human resource development. Training and Development Journal, 37 (8), 48-51.

Hays, W. L. (1963). Statistics. New York: Holt, Rinehart and Winston.

Kelley, A. I., Orgel, R. F., & Baer, D. M. (1984). Evaluation: The bottom line is closer than you think. Training and Development Journal, 38 (8), 32-37.

Kirkpatrick, D. (1975). Techniques for evaluating training programs. In D. L. Kirkpatrick (Ed.), Evaluating Training Programs (pp. 1-17). Madison, WI: American Society for Training and Development.

Kirkpatrick, D. (1977). Evaluating training problems: Evidence vs. proof. Training and Development Journal, 31 (11), 9-12.

Kirkpatrick, D. (1979). Techniques for evaluating training programs. Training and Development Journal. 33 (6), 77-92.

Kirkpatrick, D. (1983a). A practical guide for supervisory training and development. Reading, MA: Addison-Wesley.

Kirkpatrick, D. (1983b). Four steps to measuring training effectiveness. Personnel Administration, 28 (11), 19-25.

Mangum, S. L. (1984). Some evidence on criteria for choosing among alternative training techniques. Journal of Vocational Education Resource, 9 (2), 49-57.

Mosier, N. R. (1986). Financial analysis: A review of the methods and their application to employee training. St. Paul: University of Minnesota, Department of Vocational and Technical Education.

Salinger, R. D. (1981). Six reasons why training fails. In R. Zemke, L. Standke, & P. Jones (Eds.), Designing and Delivering Cost-Effective Training - and Measuring the Results (pp. 293-296). Minneapolis: Lakewood Publications.

Smeltzer, L. R. (1979). Do you really evaluate, or just talk about it? Training, 17 (8), 6-8.

Smith, M. E. (1980). Evaluating training operations and programs. Training and Development Journal, 34 (10), 70-78.

Trapnell, G. (1984). Putting the evaluation puzzle together. Training and Development Journal, 38 (5), 90-93.

Zenger, J. H. & Hargis, K. (1982). Assessing training results: It's time to take the plunge! Training and Development Journal, 36, (1), 11-16.